

# Modelling the Autocovariance of the Power Spectrum of a Solar-Type Oscillator

T. L. Campante<sup>1,2\*</sup>, C. Karoff<sup>2,3</sup>, W. J. Chaplin<sup>3</sup>, Y. P. Elsworth<sup>3</sup>, R. Handberg<sup>2</sup> and S. Hekker<sup>3</sup>

<sup>1</sup>*Centro de Astrofísica, Faculdade de Ciências, Universidade do Porto, Rua das Estrelas, 4150-762 Porto, Portugal*

<sup>2</sup>*Danish AsteroSeismology Centre, Department of Physics and Astronomy, University of Aarhus, 8000 Aarhus C, Denmark*

<sup>3</sup>*School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*

7 June 2010

## ABSTRACT

Asteroseismology is able to conduct studies on the interiors of solar-type stars from the analysis of stellar acoustic spectra. However, such an analysis process often has to rely upon subjective choices made throughout. A recurring problem is to determine whether a signal in the acoustic spectrum originates from a radial or a dipolar oscillation mode. In order to overcome this problem, we present a procedure for modelling and fitting the autocovariance of the power spectrum which can be used to obtain global seismic parameters of solar-type stars, doing so in an automated fashion without the need to make subjective choices. From the set of retrievable global seismic parameters we emphasize the mean small frequency separation and, depending on the intrinsic characteristics of the power spectrum, the mean rotational frequency splitting. Since this procedure is automated, it can serve as a useful tool in the analysis of the more than one thousand solar-type stars expected to be observed as part of the *Kepler* Asteroseismic Investigation (KAI). We apply the aforementioned procedure to simulations of the Sun. Assuming different apparent magnitudes, we address the issues of how accurately and how precisely we can retrieve the several global seismic parameters were the Sun to be observed as part of the KAI.

**Key words:** methods: data analysis – methods: statistical – stars: oscillations.

## 1 INTRODUCTION

Seismology of solar-type oscillators is a powerful tool that can be used to increase our understanding of stellar structure and evolution. Oscillations in main-sequence stars and subgiants have been measured thanks to data collected from ground-based high-precision spectroscopy (for a review see e.g., Bedding & Kjeldsen 2008) and, more recently, to photometric space-based missions such as *CoRoT* (see e.g., Appourchaux et al. 2008; Michel et al. 2008). The *Kepler* mission (for a discussion on the expected results of the asteroseismic investigation see Christensen-Dalsgaard et al. 2008; Karoff et al. 2009) will lead to a revolution in the field of asteroseismology of solar-type oscillators, since it will increase by more than two orders of magnitude the number of stars for which high-quality observations will be available, while allowing for long-term follow-ups of a selection of these targets. As of the time of writing of this article, first results arising from the *Kepler* asteroseismic programme had already been made available (Bedding et al. 2010;

Chaplin et al. 2010; Gilliland et al. 2010; Grigahcène et al. 2010; Hekker et al. 2010; Kolenberg et al. 2010; Stello et al. 2010).

Due to the large number of stars observed with *Kepler*, automated and innovative analysis pipelines/tools are needed in order to cope with the plenitude of available data (see e.g., Hekker et al. 2009; Huber et al. 2009; Kallinger et al. 2010; Karoff et al. 2010; Mathur et al. 2010; Mosser & Appourchaux 2009; Roxburgh 2009). The automated pipelines dedicated to the analysis of acoustic spectra that have been developed so far aim mainly at measuring the frequency of maximum amplitude,  $\nu_{\max}$ , the maximum mode amplitude,  $A_{\max}$ , and the mean large frequency separation,  $\Delta\nu$ . In the present study we give continuity to this work by presenting a tool capable of modelling and fitting the ACPS<sup>1</sup> (AutoCovariance of the Power Spectrum) of a solar-type oscillator. The current version of the tool accepts as free model parameters the mean small frequency separation,

\* E-mail: campante@astro.up.pt; campante@phys.au.dk

<sup>1</sup> To be precise, we compute the autocovariance of the power density spectrum.

ration parameter,  $D_0$  (see §3.1.2 for the definition of this parameter), the mean rotational frequency splitting,  $\nu_s$ , the mean linewidth,  $\langle \Gamma \rangle$ , and the stellar inclination angle,  $i$ . The output generated by this tool will hopefully contribute to further explore the diagnostic potential of solar-like oscillations, especially that of the large and small separations (see e.g., Christensen-Dalsgaard 2004). These two quantities can in turn be used to provide estimates of the radii, masses and ages of solar-like stars with consistent uncertainties (Karoff et al. 2010).

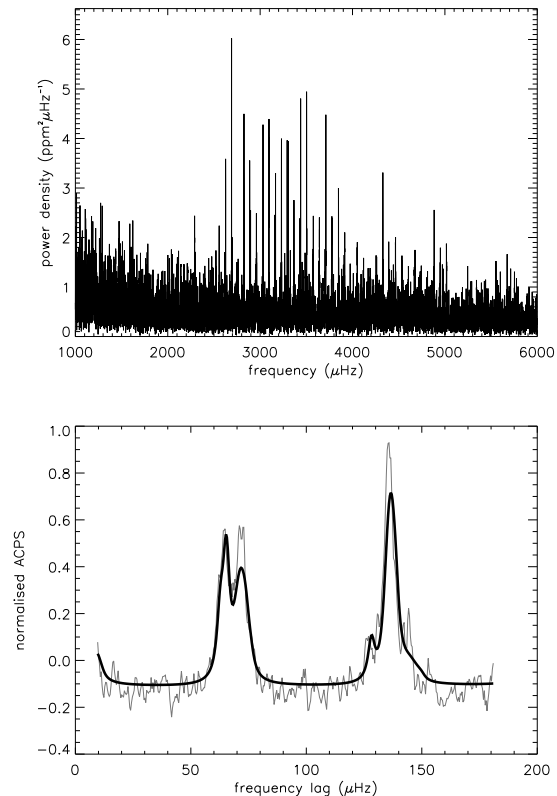
We start in §2 by providing a logical basis that led us to the development of this particular analysis tool. This is followed by a thorough description of its implementation in §3. In §4 we apply the tool to simulated data, more specifically, to a solar analog as it would be seen by *Kepler*. A discussion and conclusions are presented in §5.

## 2 RATIONALE BEHIND THE ACPS-MODELLING PROCEDURE

Unlike a canonical *peak-bagging* procedure (see e.g., Anderson, Duvall Jr. & Jefferies 1990), in the framework of the ACPS-modelling procedure individual modes do not need to be tagged by angular degree, i.e. by wave number  $\ell$ . This is required in peak-bagging to ensure the correct model is fitted to the observed modes. Instead we build a model that describes both the global and average properties of the most prominent modes, which are in turn encoded on a small set of free model parameters. This might be regarded as a major advantage, especially if we recall the ambiguity found in mode (angular degree) identification concerning the *CoRoT* target HD 49933, only recently solved after a new longer time series was made available (Appourchaux et al. 2008; Benomar et al. 2009).

As it is currently implemented, the ACPS-modelling procedure assumes the small frequency separation to be constant and not a function of frequency. Especially for evolved stars this was thought not to be a good approximation (Soriano & Vauclair 2008). However, in the light of the first *Kepler* results on red giants (Bedding et al. 2010), it is perfectly valid to assume – at least for low-luminosity red giants – a mean value of the small frequency separation between adjacent modes with  $\ell = 0$  and  $\ell = 2$ ,  $\delta\nu_{02}$ , it being in fact a nearly constant fraction of  $\Delta\nu$ . Therefore, the procedure presented herein can in principle also be employed in the case of evolved stars displaying solar-like oscillations. The ACPS-modelling procedure will however not provide sensible output when mixed modes (see e.g., Aizenman, Smeyers & Weigert 1977) are present. In any event, the presence of mixed modes will manifest clearly in the ACPS, either as a significant broadening of the peaks or by the appearance of extra peaks.

The way information is presented in the ACPS makes it very amenable for fitting. In fact, the ACPS of a solar-type oscillator will show prominent features at multiples of half the large frequency separation, a clear manifestation of the regular frequency structure of the acoustic spectrum. The shape of these features will depend upon the  $\delta\nu_{02}$  and  $\delta\nu_{01}$  spacings (see §3.1.2 for a definition of the latter), mode lifetime, rotation and stellar inclination. Fig. 1 displays an artificial acoustic spectrum of a solar analog (top panel) as



**Figure 1.** *Top Panel:* Artificial acoustic spectrum of a solar analog evidencing a regular frequency structure. *Bottom Panel:* Fit (black solid line) to the corresponding ACPS obtained using the MAP as our choice of summary statistic to represent the parameter posterior distributions. In order to enhance the visibility of the features in the ACPS we have slightly smoothed it by running a boxcar average (this is done both to the observed and to the model ACPS). Notice that the artificial star (Boris) has  $\Delta\nu \simeq 135 \mu\text{Hz}$  (see discussion in §4).

it would be obtained from a 30-day long *Kepler*’s time series together with a fit to its ACPS (bottom panel). Notice the prominent features in the ACPS around  $135 \mu\text{Hz}$  and half that value. The latter feature is clearly split into two peaks which constitutes a signature of  $\delta\nu_{01}$ . A signature of the  $\delta\nu_{02}$  spacing is also present in the wings of the feature at  $\Delta\nu$ . The widths of either feature are generally a signature of the mode lifetime of the oscillations, whereas the presence of any fine structure is a signature of rotation and stellar inclination. The ACPS depicted in Fig. 1 does indeed provide us with information about the large and small frequency separations, and mode lifetime. The question we address first is thus: *How do we extract this information from the autocovariance of the power spectrum?*

## 3 THE ACPS-MODELLING PROCEDURE

We intend to fit the observed ACPS to an ACPS created from a model of the p-mode power density spectrum (PDS) that is described by a slightly modified asymptotic relation, which in turn depends only on a few free parameters (§3.1). Although we do not fully explore its potential, we opt for a

Bayesian approach to the fitting problem for the simple reason that any prior information might be taken into account, particularly useful when constraining the parameter space (§3.2). Most importantly, we aim at obtaining a full joint posterior probability density function (PDF) concerning the set of model parameters and hence we end up employing a Markov chain Monte Carlo (MCMC) sampler (§3.3).

### 3.1 Modelling the observed ACPS

The ACPS is computed as a function of the frequency lag,  $L$ , according to the following expression:

$$\text{ACPS}(L) = \frac{1}{N} \sum_{j=0}^{N-L-1} (P_j - \bar{P})(P_{j+L} - \bar{P}), \quad (1)$$

where  $P_j$  is the particular value taken by the PDS at the fixed frequency bin  $j$ ,  $\bar{P}$  is the mean value of the sample population  $\{P_j\}$ , i.e. the mean value of the PDS within the frequency interval of interest, and  $N$  is the total number of bins within this same frequency interval. The frequency interval of interest should be in accordance with the p-mode range. The lag,  $L$ , which is expressed in terms of number of frequency bins, is allowed to vary between the corresponding frequency values of  $10 \mu\text{Hz}$  and  $4/3 \Delta\nu$ . The reason why we constrain  $L$  to this range has to do with the fact that we are primarily interested in measuring the small frequency separation. The signature of the small frequency separation is in fact mainly conveyed by the features at  $\Delta\nu/2$  and  $\Delta\nu$ . For  $L$  less than  $10 \mu\text{Hz}$  the ACPS is dominated by the signature of the window function, whereas for  $L$  greater than approximately  $4/3 \Delta\nu$  it reproduces the features seen below this value of the lag and thus becoming redundant. Finally, the ACPS is normalised after division by its maximum value.

As already stated, modelling the observed ACPS consists of two steps: (i) building a model of the subjacent p-mode PDS and (ii) computing its autocovariance as explained above. In what follows we provide the reader with the basic theory behind the power spectrum of solar-like oscillations and end up describing how the model p-mode spectrum is generated.

#### 3.1.1 The model

Solar-like oscillations are global standing acoustic waves, also known as p modes, driven by near-surface turbulent convection (see e.g., Balmforth 1992a,b). Radial oscillation modes are characterised by the radial order  $n$ , whereas non-radial modes are additionally characterised by the non-radial wave numbers  $\ell$  and  $m$ .

Ignoring any departure from spherical symmetry, non-radial modes differing only on the azimuthal wave number  $m$  are degenerate. Stellar rotation removes the  $(2\ell+1)$ -fold degeneracy of the frequency of oscillation of non-radial modes. When the angular velocity of the star,  $\Omega$ , is small and in the case of rigid-body rotation, the frequency of a  $(n, \ell, m)$  mode is given to first order by (Ledoux 1951):

$$\nu_{n\ell m} = \nu_{n\ell} + m \frac{\Omega}{2\pi} (1 - C_{n\ell}). \quad (2)$$

The kinematic splitting,  $m\Omega/(2\pi)$ , is corrected for the effect of the Coriolis force through the dimensionless quan-

tity  $C_{n\ell} > 0$ . In the asymptotic regime, i.e. for high-order, low-degree p modes, rotational splitting is dominated by advection and the splitting between adjacent modes within a multiplet is  $\nu_s \simeq \Omega/(2\pi)$ . Typically,  $\nu_{s\odot} = 0.4 \mu\text{Hz}$ .

The PDS of a single mode of oscillation is distributed around a limit spectrum with an exponential probability distribution (Woodard 1984; Duvall & Harvey 1986). This limit spectrum contains the information on the physics of the mode and may be described as a standard Lorentzian profile near the resonance. It follows that the overall limit p-mode spectrum is given by (see e.g., Fletcher et al. 2006):

$$\mathcal{P}(\nu; S_{n\ell m}, \nu_{n\ell}, \nu_s, \Gamma_{n\ell m}, N_\nu) = \sum_{n=n_0}^{n_{\max}} \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} \frac{S_{n\ell m}}{1 + \left( \frac{2(\nu - \nu_{n\ell} - m\nu_s)}{\Gamma_{n\ell m}} \right)^2} + N_\nu(\nu), \quad (3)$$

where  $S_{n\ell m}$  is the mode height,  $\nu_{n\ell}$  is the central frequency of the  $(n, \ell)$  multiplet and  $\Gamma_{n\ell m}$  is the mode linewidth. In the case of solar-type oscillators and for low angular degree  $\ell$ , we can assume that  $\Gamma$  is a function of frequency alone.  $\Gamma$  is related to the mode lifetime,  $\tau$ , through  $\Gamma = (\pi\tau)^{-1}$ .  $N_\nu$  describes the background signal originating from both granulation and activity. Notice that we are assuming that a mode is uncorrelated with any other modes or with the background signal. The stellar background signal is commonly modelled as a sum of power laws describing these physical phenomena (Harvey 1985; Aigrain, Favata & Gilmore 2004):

$$N_\nu(\nu) = \sum_{k=1}^{k_{\max}} \frac{A_k^2 B_k}{1 + (2\pi B_k \nu)^{C_k}} + N, \quad (4)$$

$\{A_k\}$  and  $\{B_k\}$  being, respectively, the corresponding amplitudes and characteristic time scales, whereas the  $\{C_k\}$  are the slopes of each of the individual power laws. A flat offset,  $N$ , is needed in order to model photon shot noise.

#### 3.1.2 The mode frequencies

Before we start modelling the observed ACPS we first need to select a suitable frequency interval for the purpose of our study. This interval should coincide with the frequency range in which p modes are located – i.e. from the frequency of the fundamental ( $n=0$ ) mode,  $\nu_f$ , up to the atmospheric acoustic cut-off frequency,  $\nu_{ac}$ .

We compute an estimate of the frequency of the fundamental mode by assuming that it scales with the mean stellar density and thus with the mean large frequency separation:

$$\nu_f = \nu_{f\odot} (\Delta\nu/\Delta\nu_{\odot}), \quad (5)$$

where  $\nu_{f\odot} = 258 \mu\text{Hz}$  and  $\Delta\nu_{\odot} = 135 \mu\text{Hz}$ .

An estimate of the atmospheric acoustic cut-off frequency is calculated by assuming that it scales with the frequency of maximum amplitude (Kjeldsen & Bedding 1995):

$$\nu_{ac} = \nu_{ac\odot} (\nu_{\max}/\nu_{\max\odot}). \quad (6)$$

For the Sun one has  $\nu_{ac\odot} = 5300 \mu\text{Hz}$  and  $\nu_{\max\odot} = 3100 \mu\text{Hz}$ .

We proceed with the assignment of frequency values to modes with degree  $\ell=0, 1, 2$  (see §3.1.3 for an explanation of this upper limit on  $\ell$ ) according to the asymptotic relation (Tassoul 1980):

$$\nu_{n\ell} = \Delta\nu(n + \frac{1}{2}\ell + \varepsilon) - \ell(\ell + 1)D_0, \quad (7)$$

where the quantity  $\varepsilon$  is sensitive to the surface layers and was taken to be 1.5. Moreover,  $D_0$  is sensitive to the sound speed gradient near the core. If Eq. (7) holds exactly then it follows that<sup>2</sup>  $\delta\nu_{02}=6D_0$ ,  $\delta\nu_{13}=10D_0$  and  $\delta\nu_{01}=2D_0$ .

The automated *Kepler* pipeline (Hekker et al. 2009) used to supply input for the procedure described herein generates as output the smooth second order change in the large frequency separation as a function of  $n$ ,  $d\Delta\nu/dn$ . In cases where an estimate of  $d\Delta\nu/dn$  is available we opt for a modified version of Eq. (7), which includes a second order correction:

$$\nu_{n\ell} = \Delta\nu(n + \frac{1}{2}\ell + \varepsilon) - \ell(\ell + 1)D_0 + (n - n_{\max})^2 \frac{d\Delta\nu/dn}{2}. \quad (8)$$

The overtone with the highest power,  $n_{\max}$ , is given by<sup>3</sup>  $\text{round}[(\nu_{\max} - \nu_{\ell})/\Delta\nu]$ .

For the sake of clarity we should again stress that individual modes in the observed PDS do not need to be tagged by angular degree  $\ell$  prior to the fit. Assignment of frequency values to modes in the model of the PDS, according to either Eq. (7) or Eq. (8), is in fact done in complete ignorance of the correct mode tagging.

### 3.1.3 The mode heights

The height of a multiplet component can be expressed as:

$$S_{n\ell m} = \mathcal{E}_{\ell m}(i) S_{n\ell} = \mathcal{E}_{\ell m}(i) V_{\ell}^2 \alpha_{n\ell}. \quad (9)$$

The  $\mathcal{E}_{\ell m}(i)$  factor represents mode visibility within a multiplet and  $i$  is the inclination angle between the direction of the stellar rotation axis and the line of sight. This factor is given by (Gizon & Solanki 2003):

$$\mathcal{E}_{\ell m}(i) = \frac{(\ell - |m|)!}{(\ell + |m|)!} \left[ P_{\ell}^{|m|}(\cos i) \right]^2, \quad (10)$$

where  $P_{\ell}^m(x)$  are the associated Legendre functions. The quantity  $V_{\ell}^2$  is an estimate of the geometrical visibility of the total power in a  $(n, \ell)$  multiplet as a function of  $\ell$ , whereas  $\alpha_{n\ell}$  depends mainly on the frequency and excitation mechanism. Eq. (9) is only strictly valid under one assumption: When the stellar flux is integrated over the full apparent disc, one must assume that the weighting function depends only on the distance to the disc centre. In this case, the apparent mode amplitude can effectively be separated into two factors:  $\mathcal{E}_{\ell m}(i)$  and  $V_{\ell}^2$ . This assumption holds very well in the case of intensity measurements, since the weighting function is then mainly linked to the limb-darkening, whereas for velocity measurements, departures might be observed due to asymmetries in the velocity field induced by rotation (see Ballot, García & Lambert 2006; Ballot et al. 2008, and references therein). The heights of non-radial modes are commonly defined based on the heights of radial modes according to Eq. (9), and taking into account the  $V_{\ell}/V_0$  ratios (see e.g., Bedding et al. 1996). Note that  $\ell=0$  modes constitute a sensible reference since they are not split by rotation.

<sup>2</sup>  $\delta\nu_{13}$  is the spacing between adjacent modes with  $\ell=1$  and  $\ell=3$ , and  $\delta\nu_{01}$  is the amount by which  $\ell=1$  modes are offset from the midpoint between the  $\ell=0$  modes on either side.

<sup>3</sup> The  $\text{round}[\ ]$  function rounds the argument to its closest integer.

In the present case of stellar observations or when observing the Sun-as-a-star, the associated whole-disc light integration strongly suppresses high-degree modes due to the lack of spatial resolution. Stellar observations are hence mostly sensitive to high-order acoustic eigenmodes with  $\ell \leq 3$ .

We compute the power envelope for radial modes as a function of frequency according to Kjeldsen et al. (2008). We start by subtracting a fit to the background signal from the observed PDS. The residual spectrum thus obtained is heavily smoothed over the range occupied by the p modes by convolving it with a Gaussian having a FWHM of  $4\Delta\nu$ . Finally, we multiply the smoothed, residual spectrum by  $\Delta\nu/c$ , where  $c$  is a factor that measures the effective number of modes per order and taken to be 3.03. The height,  $S_{n0}$ , of a radial mode in units of power density is then given by (Chaplin et al. 2008):

$$S_{n0} = \frac{2A^2 T}{\pi T T_{n0} + 2}, \quad (11)$$

where  $A^2$  is the total power of the mode (as determined from the power envelope for radial modes) and  $T$  is the effective observational length.

### 3.1.4 Setting up the model

In order to generate the model p-mode spectrum, our modelling procedure requires the following input: The observed PDS sampled at the true resolution (i.e. computed over a grid of uncorrelated frequency bins), a fit to the background signal,  $\nu_{\max}$  and  $\Delta\nu$ . The gradient of the large frequency separation with  $n$ ,  $d\Delta\nu/dn$ , is optional. This input, with the exception of the observed PDS, might be obtained from the automated *Kepler* pipelines.

A common linewidth value is assigned to all the modes (a sensible assumption if the range in frequency of the spectrum being tested is not too wide). The fitted parameter  $\langle\Gamma\rangle$  might then be interpreted as a mean linewidth.

The model p-mode spectrum (including the background signal) is finally assembled according to Eq. (3), taking into account the global seismic parameters fitted in our procedure, namely,  $D_0$ ,  $\nu_s$ ,  $\langle\Gamma\rangle$  and  $i$ . In case we would want to allow for the effect of the windowing and statistical weighting, then we should at this stage convolve the model p-mode spectrum with the spectral window.

## 3.2 Bayesian inference

### 3.2.1 Parameter estimation

In the framework of Bayesian parameter estimation, a particular model of the observed ACPS is assumed to be true and the hypothesis space of interest relates to the values taken by the model parameters,  $\Theta = \{D_0, \nu_s, \langle\Gamma\rangle, i\}$ . These parameters are continuous, meaning that the quantity of interest is a PDF, in contrast to traditional point estimate methods. Let us state Bayes' theorem:

$$p(\Theta|D, I) = \frac{p(\Theta|I) p(D|\Theta, I)}{p(D|I)}, \quad (12)$$

where  $D$  represents the available data and the prior information is represented by  $I$ .  $p(\Theta|I)$  is called the *prior probability*, whereas  $p(\Theta|D, I)$  is called the *posterior probability*.

The *likelihood function* is represented by  $p(D|\Theta, I)$ .  $p(D|I)$ , the *global likelihood*, simply plays the role of a normalisation constant.

The procedure of *marginalisation* makes it possible to derive the marginal posterior PDF for a subset of parameters  $\Theta_A$ , by integrating out the remaining parameters  $\Theta_B$ , the so-called *nuisance parameters*:

$$p(\Theta_A|D, I) = \int p(\Theta_A, \Theta_B|D, I) d\Theta_B. \quad (13)$$

Furthermore, assuming that the prior on  $\Theta_A$  is independent of the prior on the remaining parameters, then by applying the product rule we have:

$$p(\Theta_A, \Theta_B|I) = p(\Theta_A|I)p(\Theta_B|\Theta_A, I) = p(\Theta_A|I)p(\Theta_B|I). \quad (14)$$

The main advantage of the Bayesian framework when compared to a frequentist approach is the ability to incorporate relevant prior information through Bayes' theorem and to evaluate its effect on our analysis (see e.g., Brewer 2008, for possible applications of Bayesian probability theory in Astrophysics).

### 3.2.2 The likelihood function

We would like to specify the likelihood function, i.e. the joint PDF for the data sample. We know that, for a given frequency lag  $L_i \in [L_{\min}, L_{\max}]$ :

$$\text{ACPS}_o(L_i) = \text{ACPS}_m(L_i) + e_i, \quad (15)$$

where  $\text{ACPS}_o$  and  $\text{ACPS}_m$  are the observed and model ACPS, respectively. The error term  $e_i$  follows a Gaussian<sup>4</sup> distribution with zero mean and standard deviation  $\sigma_i$ . Hence, assuming that the model is deterministic, i.e. true, we can write

$$f[\text{ACPS}_o(L_i)] = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{e_i^2}{2\sigma^2}\right\} = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{[\text{ACPS}_o(L_i) - \text{ACPS}_m(L_i)]^2}{2\sigma^2}\right\}, \quad (16)$$

where  $f[\text{ACPS}_o(L_i)]$  is the probability density that the observed ACPS takes a particular value  $\text{ACPS}_o(L_i)$  at a frequency lag  $L_i$ . Notice that we are assuming  $\sigma_i$  to be constant over the whole range spanned by  $L_i$ .

By ignoring the effect of correlation between points in the ACPS we arrive at the following expression for the likelihood function:

$$\mathcal{L}(\Theta) \equiv p(D|\Theta, I) = \prod_i f[\text{ACPS}_o(L_i)]. \quad (17)$$

While the effect of correlation clearly must be present, the fact that we ignore it can be shown to only affect attempted error calculations and not the fitted values themselves (see Fletcher et al. 2006, and references therein). This does not constitute a problem in the present case, since we will be employing a MCMC sampler in order to obtain the marginal posteriors for each of the model parameters, from which we can estimate the uncertainties.

<sup>4</sup> Each point in the ACPS is the sum of a large number of independent and identically distributed random variables. Therefore, the Central Limit Theorem applies and the distribution of  $e_i$  should approach the normal distribution.

### 3.3 Markov chain Monte Carlo (MCMC)

After inspection of Eq. (13), the need for a mathematical tool that is able to efficiently evaluate the multi-dimensional integrals required in the computation of the marginal posteriors becomes clear. This constitutes the principal rationale behind the method known as *Markov chain Monte Carlo*, which aims at drawing samples from the *target distribution*,  $p(\Theta|D, I)$ , by constructing a pseudo-random walk in model parameter space. Such a pseudo-random walk is achieved by generating a Markov chain, whereby a new sample,  $\Theta_{t+1}$ , depends solely on the previous sample,  $\Theta_t$ , in accordance with a time-independent quantity called the *transition kernel*,  $p(\Theta_{t+1}|\Theta_t)$ :

$$p(\Theta_{t+1}|\Theta_t) = q(\Theta_{t+1}|\Theta_t) \min\left[1, \frac{p(\Theta_{t+1}|D, I)}{p(\Theta_t|D, I)} \frac{q(\Theta_t|\Theta_{t+1})}{q(\Theta_{t+1}|\Theta_t)}\right], \quad (18)$$

where  $q(\Theta_{t+1}|\Theta_t)$  is a *proposal distribution* centred on  $\Theta_t$ . After a *burn-in* phase,  $p(\Theta_{t+1}|\Theta_t)$  is able to generate samples of  $\Theta$  with a probability density converging on the target distribution. The algorithm that we employ in order to generate a Markov chain was initially proposed by Metropolis et al. (1953), and subsequently generalised by Hastings (1970), this latter version being commonly referred to as the *Metropolis–Hastings algorithm*.

The Metropolis–Hastings MCMC algorithm just outlined above runs the risk of becoming stuck in a local maximum of the target distribution, thus failing to fully explore all regions in parameter space containing significant probability. A way of overcoming this difficulty is to employ *parallel tempering* (see e.g., Earl & Deem 2005), whereby a discrete set of progressively flatter versions of the target distribution is created by introducing a *temperature parameter*,  $\mathcal{T}$ . In practice, use is made of its reciprocal,  $\beta = 1/\mathcal{T}$ , referred to as the *tempering parameter*. By modifying Eq. (12), we generate the tempered distributions as follows:

$$p(\Theta|D, \beta, I) = C p(\Theta|I) p(D|\Theta, I)^\beta, \quad 0 < \beta \leq 1, \quad (19)$$

where  $C$  is a constant. For  $\beta = 1$ , we retrieve the target distribution, also called the *cold sampler*, whereas for  $\beta < 1$ , the hotter distributions are effectively flatter versions of the target distribution. By running such a set of cooperative chains in parallel and by further allowing for the swap of their respective parameter states, we enable the algorithm to sample the target distribution in a way that allows for both the investigation of its overall features (low- $\beta$  chains) and the examination of the fine details of a local maximum (high- $\beta$  chains). See Appendix A for the current version of the parallel tempering Metropolis–Hastings algorithm written in pseudocode.

Moreover, based on a statistical *control system* (CS) similar to the one described in Gregory (2005), we automate the process of calibration of the Gaussian proposal  $\sigma_{\text{CS}}$ -values, which specify the direction and step size in parameter space when proposing a new sample to be drawn. The optimal choice of  $\{\sigma_{\text{CS}}\}$  is closely related to the average rate at which proposed state changes are accepted, the so-called *acceptance rate*. The control system makes use of an error signal to steer the selection of the  $\sigma_{\text{CS}}$ -values during the burn-in stage of a single parallel tempering MCMC run, acting independently on each of the tempered chains.

The error signal is proportional to the difference between the current acceptance rate and the target acceptance rate. As soon as the error signal for each of the tempered chains is less than a measure of the statistical fluctuation expected for a zero mean error, the control system is turned off and the algorithm switches to the standard parallel tempering MCMC.

#### 4 APPLICATION TO SIMULATED DATA: THE SOLAR TWIN BORIS

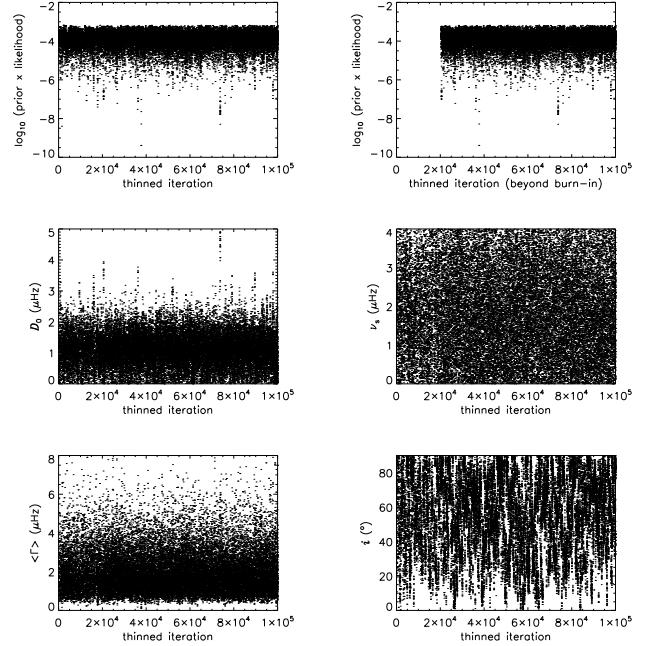
We display the results obtained when applying the ACPs-modelling procedure to the particular case of Boris, an artificial main-sequence star created in the framework of the AsteroFLAG group for the purpose of conducting hare and hounds. Briefly, Boris is what we might call a solar analog ( $L = 1.00 L_{\odot}$ ;  $T_{\text{eff}} = 5778 \text{ K}$ ;  $R = 1.00 R_{\odot}$ ). We refer the reader to Chaplin et al. (2008) and Stello et al. (2009) for further insight into AsteroFLAG’s activities.

##### 4.1 Experimental setup

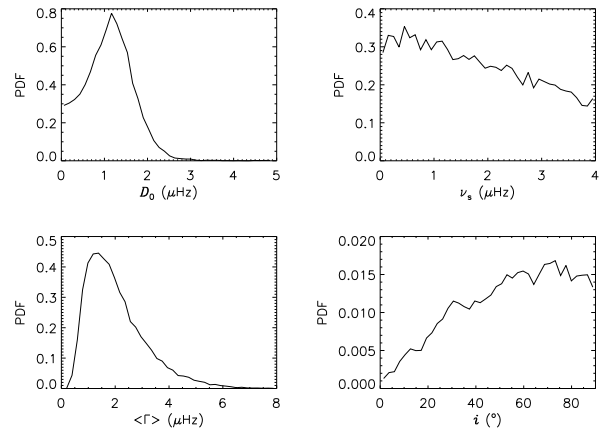
Boris was assumed to have an apparent visual magnitude of either  $V = 8$  or  $V = 9$  or  $V = 10$  (these values correspond to the bright end of the nominal apparent magnitude target range for *Kepler*). In each case, 250 realisations of a PDS were generated. The goal was to assess both the accuracy and the precision of the several Bayesian summary statistics of the fitted parameters, namely, the MAP (Maximum *A Posteriori*), the median, the mean and the marginal posterior mode. These power density spectra were directly generated in Fourier space and include contributions arising from p modes, granulation, activity, photon shot noise and instrumental noise. They correspond to 30-day long time series with a 60-second cadence, as it would be expected for *Kepler*’s high-cadence survey targets. The stellar inclination angle,  $i$ , entering the simulations was allowed to vary and was drawn from a probability distribution that assumes random orientations –  $\sin(i)$ . Reference values are given for the other model parameters by:  $D_0^{\text{ref}} = 1.43 \mu\text{Hz}$ ,  $\nu_s^{\text{ref}} = \nu_{s\odot} = 0.4 \mu\text{Hz}$  and  $\langle \Gamma \rangle^{\text{ref}} = 1.95 \mu\text{Hz}$ . All the remaining relevant quantities that enter the simulations were assigned ‘solar’ values.

A prototype of the Birmingham–Sheffield Hallam automated pipeline (Hekker et al. 2009) was then run on the simulated power density spectra in order to retrieve useful input for the ACPs-modelling procedure, to be specific, a fit to the background signal,  $\Delta\nu$  and  $d\Delta\nu/dn$ . The frequency interval of interest (for purposes of the computation of the ACPs) ranges from 2100 to 4500  $\mu\text{Hz}$  and  $\nu_{\text{max}}$  has been fixed at 3300  $\mu\text{Hz}$ .

Figs. 1, 2 and 3 display the typical graphical output of the ACPs-modelling procedure resulting from the analysis of a single realisation of the PDS of Boris (with  $V = 8$  and a random orientation characterised by  $i = 32.4^\circ$ ). Note that throughout the analysis of the full set of realisations of the power spectrum, uniform priors have been imposed on  $D_0$ ,  $\nu_s$  and  $\langle \Gamma \rangle$ , whereas  $p(i|I) = \sin(i)$ .



**Figure 2.** Sequence of  $1 \times 10^5$  samples (at a thinning interval of 1) drawn from the target distribution ( $\beta = 1$ ) by a Metropolis–Hastings Markov chain Monte Carlo. *Top Panels:* Behaviour of  $\log_{10}[p(\Theta|I)\mathcal{L}(\Theta)]$  as a function of the iteration number (a means of visually inspecting the convergence of the MCMC). *Middle and Bottom Panels:* Behaviour of model parameter values as a function of the iteration number.



**Figure 3.** Binned marginal posterior PDFs of the model parameters. These are built after assembling the corresponding samples (beyond burn-in) depicted in the middle and bottom panels of Fig. 2. Based on these marginal posteriors we may then evaluate the several Bayesian summary statistics for each of the model parameters.

##### 4.2 On retrieving $D_0$ and $\langle \Gamma \rangle$

Figs. 4 and 5 result from the application of the ACPs-modelling procedure to the full set of realisations of the power spectrum (250 realisations for each value of  $V$ ), allowing one to assess the potential biases of the Bayesian summary statistics of the fitted parameters  $D_0$  and  $\langle \Gamma \rangle$ , re-

spectively. The precision associated with these same estimators is also indicated in the plots.

When choosing a summary statistic to be used, we should make sure that it is representative of the marginal probability distribution of the parameter in question, a point that all the summary statistics considered seem to satisfy. The marginal posterior PDFs of  $D_0$  and  $\langle \Gamma \rangle$  are asymmetric (see Fig. 3) and well modelled by (Fierry Fraillon et al. 1998):

$$f(x) = K(x/s)^\alpha \exp[(x/s)^{\alpha+1}], \quad (20)$$

where  $K$  is a normalisation constant,  $s$  is an adjustable parameter describing the shape of the distribution and  $\alpha$  defines the type of distribution ( $\alpha=0$  for a Boltzmann law and high values of  $\alpha$  for a Gaussian law). Furthermore, the set of summary parameter values should provide a good fit to the data, the quality of which might be assessed by computing the rms of the residuals. This latter criterion led us to adopt the MAP as our choice of summary statistic.

It turns out from Figs. 4 and 5 that the different estimators, although equally robust, present biases whose magnitude and sign vary with the estimator being considered, a direct consequence of the asymmetry of the marginal posterior PDFs of  $D_0$  and  $\langle \Gamma \rangle$ . A way of overcoming these inherent biases would be to present these values normalised to the homologous values obtained after performing a similar analysis on Sun-as-a-star data covering an equivalent range in frequency (i.e. scaling by the ratio of the acoustic cut-off frequencies of the Sun and the target in question).

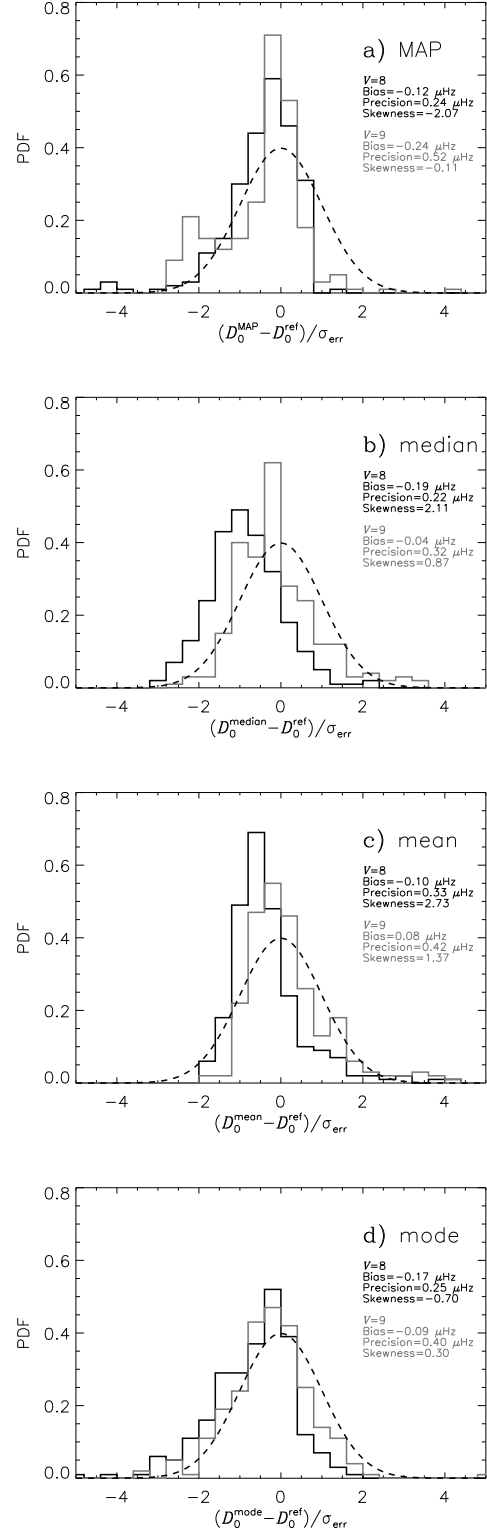
The MAP summary statistic allows one to determine  $D_0$  with a relative error of 8 per cent and 17 per cent for  $V=8$  and  $V=9$ , respectively. On the other hand, determination of  $\langle \Gamma \rangle$  is accomplished with a relative error of 13 per cent and 18 per cent for  $V=8$  and  $V=9$ , respectively.

A clear degradation of the precision associated with the estimators is seen when considering the set of fainter objects ( $V=9$ ). This is due to the fact that  $V$  is directly coupled to the photon shot noise level and hence to  $S/N$ .

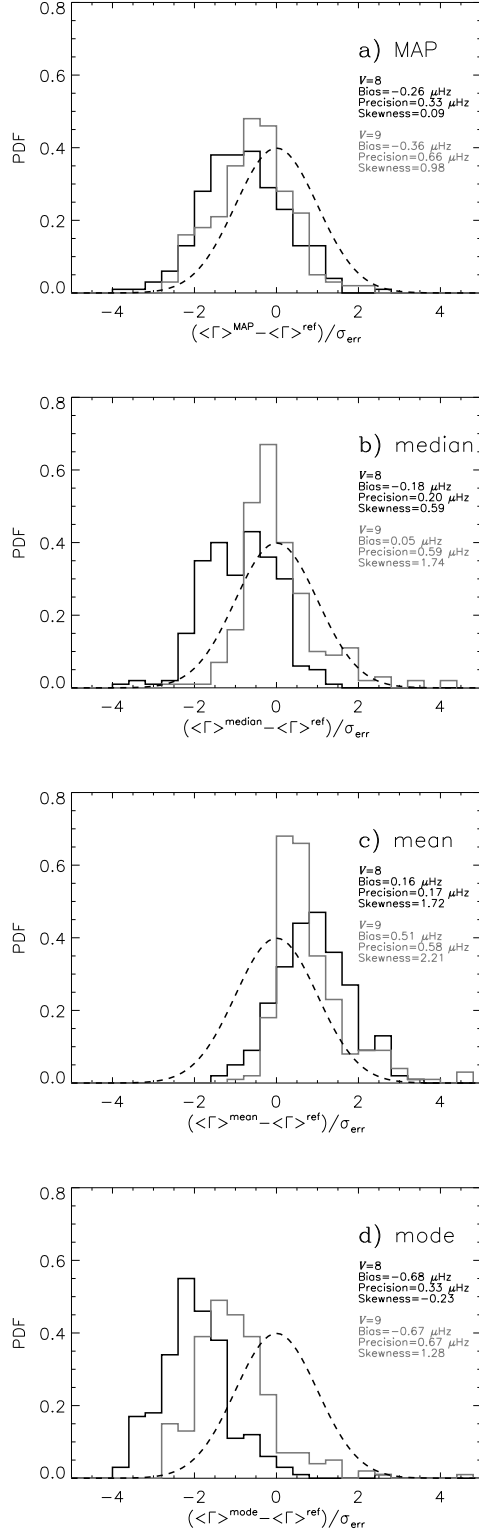
Notice that no results have been plotted concerning the set characterised by  $V=10$  as no sensible output was generated by the ACPS-modelling procedure. We should mention that at this  $S/N$ , the automated pipeline is no longer capable of supplying an input value for  $d\Delta\nu/dn$  and hence the model p-mode spectrum generated by the ACPS-modelling procedure assumes a constant large separation with  $n$ .

#### 4.3 On retrieving $i$ and $\nu_s$

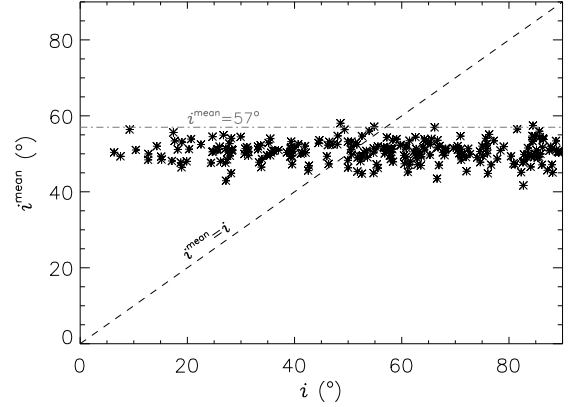
With the current experimental setup, we face a scenario where the linewidths of the individual modes present in the simulated power spectra are systematically larger than the rotational splitting, i.e.  $\nu_s \lesssim \Gamma$ . Consequently, multiplet components are blended together, which strongly correlates the inclination with the rotational splitting, making a successful retrieval of these global parameters not feasible. As a further matter, the spectral resolution is too low ( $\delta\nu = 0.39 \mu\text{Hz}$ ) and will only marginally allow the rotational splitting to be resolved. By again looking at the PDF of the inclination for a single realisation of the power spectrum in Fig. 3, we realise that we are in fact retrieving the imposed prior since there is no enough evidence in the data to proceed otherwise. Fig. 6 corroborates this, if we bear



**Figure 4.** PDFs of the actual errors in the determination of  $D_0$ , which in turn have been normalised after division by their standard deviation,  $\sigma_{\text{err}}$ . Several Bayesian summary statistics have been considered: a) the MAP, b) the median, c) the mean and d) the mode. The solid black line corresponds to the set with  $V=8$ , whilst the solid grey line corresponds to the set with  $V=9$ . The dashed line represents the standard normal distribution.



**Figure 5.** The same as in Fig. 4 but now regarding the estimation of  $\langle \Gamma \rangle$ .



**Figure 6.** The posterior mean of the stellar inclination angle,  $i^{\text{mean}}$ , versus the input inclination,  $i$ . This plot refers to the set of simulations with  $V=8$ .

in mind that  $i^{\text{mean}} \equiv \int_0^{90^\circ} i p(i|D, I) di \simeq \int_0^{90^\circ} i p(i|I) di = \int_0^{90^\circ} i \sin(i) di = 57^\circ$ . Similarly, by looking at the PDF of the rotational splitting for a single realisation of the power spectrum in Fig. 3, we notice that we are in fact retrieving the imposed uniform prior, although slightly biased toward low values of the splitting.

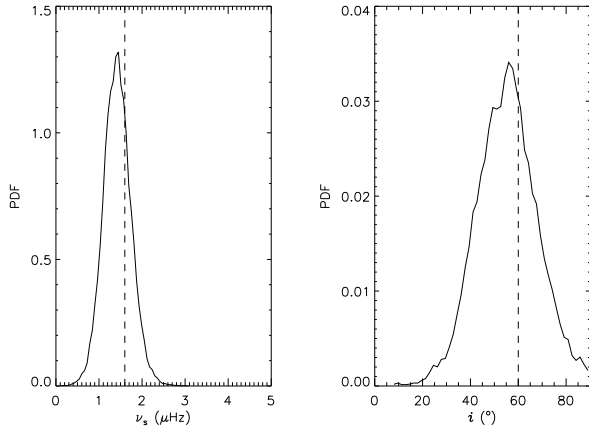
In order to demonstrate the full capability of this tool, we have applied it to a series of higher  $S/N$  realisations of the PDS of Boris (assumed to have  $V=8$ ), which now correspond to 6-month long time series ( $\delta\nu = 0.06 \mu\text{Hz}$ ). Since we wish to avoid a scenario where  $\nu_s \lesssim \Gamma$  and enter a more favourable regime where instead  $\Gamma < \nu_s \lesssim \delta\nu_{02}$ , we (i) increased the reference value for the mean rotational splitting,  $\nu_s^{\text{ref}} \in \{2\nu_{s\odot}, 4\nu_{s\odot}, 6\nu_{s\odot}\}$ , while (ii) simultaneously adopting a narrower frequency interval of interest, ranging from 2100 to 3300  $\mu\text{Hz}$ , at the low-frequency end of the acoustic spectrum. This latter step will effectively decrease the reference value for the mean linewidth,  $\langle \Gamma \rangle^{\text{ref}}$ , since we use solar mode linewidths when generating the artificial power density spectra, which are known to increase toward higher frequencies. Finally, we adopted  $i^{\text{ref}} = 60^\circ$  and imposed uniform priors on all the model parameters. Fig. 7 illustrates a case where the inclination and the rotational splitting have been successfully retrieved, after applying the ACPS-modelling procedure to a single realisation of the PDS of Boris.

## 5 SUMMARY AND DISCUSSION

(i) We have developed a new data analysis tool based on modelling and fitting the autocovariance of the acoustic PDS of a solar-type oscillator. Its main advantage when compared to a canonical peak-bagging procedure, relies on the fact that there is no need to carry out mode (angular degree  $\ell$ ) identification prior to performing the fit. This procedure is in principle also suitable for being employed in the case of evolved stars displaying solar-like oscillations.

(ii) The implementation of the ACPS-modelling procedure has been thoroughly described. Furthermore, its automated character makes this procedure appropriate for the





**Figure 7.** PDFs of  $\nu_s$  and  $i$  resulting from the application of the ACPS-modelling procedure to a single realisation of the PDS of Boris. Reference values are given by  $\nu_s^{\text{ref}} = 4 \nu_{s\odot} = 1.6 \mu\text{Hz}$  and  $i^{\text{ref}} = 60^\circ$ , and are represented by dashed vertical lines. MAP estimates are in turn given by  $\nu_s^{\text{MAP}} = 1.44^{+0.26}_{-0.36} \mu\text{Hz}$  and  $i^{\text{MAP}} = 55.9^\circ \pm 11.8^\circ$ , therefore being within uncertainties of the reference values.

analysis of a large number of data sets (e.g. arising from the *Kepler* mission).

(iii) The ACPS conveys information on the large and small frequency separations, mode lifetime and rotation. The current version of the ACPS-modelling procedure accepts  $D_0$ ,  $\nu_s$ ,  $\langle\Gamma\rangle$  and  $i$  as free parameters. The prospective inclusion of  $\Delta\nu$  and  $d\Delta\nu/dn$  as additional free parameters is envisaged.

(iv) The tool has been applied to simulated data mimicking what would have been expected for a solar analog observed at high cadence during *Kepler*'s survey phase. We assessed the potential biases as well as the precision associated with the several Bayesian summary statistics of the fitted parameters  $D_0$  and  $\langle\Gamma\rangle$ , having been able to decide on a preferred summary statistic, namely, the MAP. We addressed a way of overcoming the inherent biases. These biases would not in any case underpin the usefulness of this procedure, since the preferred summary statistic is given together with robust estimates of the uncertainties: a MCMC sampler is employed in order to obtain the marginal posteriors for each of the model parameters, from which we can estimate the uncertainties. For instance, using the SEEK routine (Quirion et al., in preparation) in order to estimate stellar parameters and basing it on a set of asteroseismic parameters returned by this procedure, would lead to the estimation of a set of stellar parameters possessing robust uncertainties (see discussion in Karoff et al. 2010).

(v) No sensible output has been generated by the ACPS-modelling procedure for the set with  $V = 10$ . We could argue, of course, that by increasing the effective observational length of the simulated data, the method could successfully reach a lower  $S/N$  since realisation noise is expected to scale as  $1/\sqrt{T}$ .

(vi) Furthermore, retrieval of both  $i$  and  $\nu_s$  could not be achieved with the current experimental setup and a reason for that was given based on the intrinsic characteristics of the simulated spectra. An example has however been in-

cluded whereby the full capability of this tool is demonstrated.

## ACKNOWLEDGMENTS

TLC is supported by grant with reference number SFRH/BD/36240/2007 from FCT/MCTES, Portugal. CK acknowledges financial support from the Danish Natural Sciences Research Council. WJC, YPE and SH acknowledge the support of the Science and Technology Facilities Council (STFC), UK. We wish to thank the anonymous referee for valuable comments.

## REFERENCES

- Aigrain S., Favata F., Gilmore G., 2004, *A&A*, 414, 1139
- Aizenman M., Smeyers P., Weigert A., 1977, *A&A*, 58, 41
- Anderson E. R., Duvall Jr. T. L., Jefferies S. M., 1990, *ApJ*, 364, 699
- Appourchaux T., Michel E., Auvergne M. et al., 2008, *A&A*, 488, 705
- Ballot J., Appourchaux T., Toutain T., Guittet M., 2008, *A&A*, 486, 867
- Ballot J., García R. A., Lambert P., 2006, *MNRAS*, 369, 1281
- Balmforth N. J., 1992a, *MNRAS*, 255, 639
- Balmforth N. J., 1992b, *MNRAS*, 255, 603
- Bedding T. R., Huber D., Stello D. et al., 2010, *ApJ*, 713, L176
- Bedding T. R., Kjeldsen H., 2008, in G. van Belle ed., 14th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun Vol. 384 of Astronomical Society of the Pacific Conference Series, Asteroseismology from Solar-Like Oscillations. pp 21–+
- Bedding T. R., Kjeldsen H., Reetz J., Barbuy B., 1996, *MNRAS*, 280, 1155
- Benomar O., Baudin F., Campante T. L. et al., 2009, *A&A*, 507, L13
- Brewer B. J., 2008, ArXiv e-prints
- Chaplin W. J., Appourchaux T., Arentoft T. et al., 2008, *Astronomische Nachrichten*, 329, 549
- Chaplin W. J., Appourchaux T., Elsworth Y. et al., 2010, *ApJ*, 713, L169
- Chaplin W. J., Houdek G., Appourchaux T., Elsworth Y., New R., Toutain T., 2008, *A&A*, 485, 813
- Christensen-Dalsgaard J., 2004, *Sol. Phys.*, 220, 137
- Christensen-Dalsgaard J., Arentoft T., Brown T. M., Gilliland R. L., Kjeldsen H., Borucki W. J., Koch D., 2008, *Journal of Physics Conference Series*, 118, 012039
- Duvall Jr. T. L., Harvey J. W., 1986, in Gough D. O., ed., *NATO ASIC Proc. 169: Seismology of the Sun and the Distant Stars Solar Doppler shifts: Sources of continuous spectra*. pp 105–116
- Earl D. J., Deem M. W., 2005, *Physical Chemistry Chemical Physics (Incorporating Faraday Transactions)*, 7, 3910
- Fierry Fraillon D., Gelly B., Schmider F. X., Hill F., Fossat E., Pantel A., 1998, *A&A*, 333, 362
- Fletcher S. T., Chaplin W. J., Elsworth Y., Schou J., Buzasi D., 2006, *MNRAS*, 371, 935

- Gilliland R. L., Brown T. M., Christensen-Dalsgaard J. et al., 2010, *PASP*, 122, 131
- Gizon L., Solanki S. K., 2003, *ApJ*, 589, 1009
- Gregory P. C., 2005, *Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with ‘Mathematica’ Support*. Cambridge University Press
- Grigahcène A., Antoci V., Balona L. et al., 2010, *ApJ*, 713, L192
- Harvey J., 1985, Technical report, High-resolution helioseismology
- Hastings W. K., 1970, *Biometrika*, 57, 97
- Hekker S., Broomhall A., Chaplin W. J. et al., 2009, *MNRAS*, pp 1945–+
- Hekker S., Debosscher J., Huber D. et al., 2010, *ApJ*, 713, L187
- Huber D., Stello D., Bedding T. R., Chaplin W. J., Arentoft T., Quirion P., Kjeldsen H., 2009, *Communications in Asteroseismology*, 160, 74
- Kallinger T., Weiss W. W., Barban C. et al., 2010, *A&A*, 509, A77+
- Karoff C., Campante T. L., Chaplin W. J., 2010, *ArXiv e-prints*
- Karoff C., Metcalfe T. S., Chaplin W. J., Elsworth Y., Kjeldsen H., Arentoft T., Buzasi D., 2009, *MNRAS*, 399, 914
- Kjeldsen H., Bedding T. R., 1995, *A&A*, 293, 87
- Kjeldsen H., Bedding T. R., Arentoft T. et al., 2008, *ApJ*, 682, 1370
- Kolenberg K., Szabó R., Kurtz D. W. et al., 2010, *ApJ*, 713, L198
- Ledoux P., 1951, *ApJ*, 114, 373
- Mathur S., García R. A., Régulo C. et al., 2010, *A&A*, 511, A46+
- Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, *J. Chem. Phys.*, 21, 1087
- Michel E., Baglin A., Auvergne M. et al., 2008, *Science*, 322, 558
- Mosser B., Appourchaux T., 2009, *A&A*, 508, 877
- Roxburgh I. W., 2009, *A&A*, 506, 435
- Soriano M., Vauclair S., 2008, *A&A*, 488, 975
- Stello D., Basu S., Bruntt H. et al., 2010, *ApJ*, 713, L182
- Stello D., Chaplin W. J., Bruntt H. et al., 2009, *ApJ*, 700, 1589
- Tassoul M., 1980, *ApJS*, 43, 469
- Woodard M. F., 1984, PhD thesis, AA(UNIVERSITY OF CALIFORNIA, SAN DIEGO.)

## APPENDIX A: PARALLEL TEMPERING METROPOLIS–HASTINGS

```

1: procedure PT METROPOLIS–HASTINGS
2:    $\Theta_{0,i} = \Theta_0$ ,  $1 \leq i \leq n_\beta$ 
3:   for  $t = 0, 1, \dots, n_{\text{it}} - 1$  do
4:     for  $i = 1, 2, \dots, n_\beta$  do
5:       Propose a new sample to be drawn from a
         proposal distribution:  $\Lambda \sim N(\Theta_{t,i}; \mathbf{C}_i)$ 
6:       Compute the Metropolis ratio:
          $r = \frac{p(\Lambda|D, \beta_i, I) q(\Theta_{t,i}|\Lambda)}{p(\Theta_{t,i}|D, \beta_i, I) q(\Lambda|\Theta_{t,i})}$ 
7:       Sample a uniform random variable:
          $U_1 \sim \text{Uniform}(0, 1)$ 
8:       if  $r \geq U_1$  then
9:          $\Theta_{t+1,i} = \Lambda$ 
10:      else
11:         $\Theta_{t+1,i} = \Theta_{t,i}$ 
12:      end if
13:    end for
14:     $U_2 \sim \text{Uniform}(0, 1)$ 
15:    if  $1/n_{\text{swap}} \geq U_2$  then
16:      Select random chain:
         $i \sim \text{UniformInt}(1, n_\beta - 1)$ 
17:       $j = i + 1$ 
18:      Compute  $r_{\text{swap}}$ :
         $r_{\text{swap}} = \frac{p(\Theta_{t,j}|D, \beta_i, I) p(\Theta_{t,i}|D, \beta_j, I)}{p(\Theta_{t,i}|D, \beta_i, I) p(\Theta_{t,j}|D, \beta_j, I)}$ 
19:       $U_3 \sim \text{Uniform}(0, 1)$ 
20:      if  $r_{\text{swap}} \geq U_3$  then
21:        Swap parameter states of chains  $i$  and  $j$ :
         $\Theta_{t,i} \leftrightarrow \Theta_{t,j}$ 
22:      end if
23:    end if
24:  end for
25:  return  $\Theta_{t,i}$ ,  $\forall t, i: \beta_i = 1$ 
26: end procedure

```

**Figure A1.** Pseudocode-written version of the parallel tempering Metropolis–Hastings algorithm.  $n_\beta$  is the total number of tempered chains,  $n_{\text{it}}$  is the number of iterations when running the MCMC,  $q(\Lambda|\Theta_{t,i}) = N(\Theta_{t,i}; \mathbf{C}_i)$  is a multivariate Gaussian distribution centred on  $\Theta_{t,i}$  and with diagonal covariance matrix  $\mathbf{C}_i$ , and  $n_{\text{swap}}$  is the mean number of iterations between successive proposals to swap the parameter states of two adjacent chains.